

# Normal Factor Graphs: A Diagrammatic Approach to Linear Algebra

Ali Al-Bashabsheh  
School of Information Technology and Engineering  
University of Ottawa  
Ottawa, Canada  
{aalba059, yymao}@site.uottawa.ca

Pascal O. Vontobel  
Hewlett-Packard Laboratories  
1501 Page Mill Road  
Palo Alto, CA 94304, USA  
pascal.vontobel@ieee.org

**Abstract**—Inspired by some new advances on normal factor graphs (NFGs), we introduce NFGs as a simple and intuitive diagrammatic approach towards encoding some concepts from linear algebra. We illustrate with examples the workings of such an approach and settle a conjecture of Peterson on the Pfaffian.

## I. INTRODUCTION

The application of graphical modeling frameworks and of the computational tools therein has undoubtedly revolutionized the field of coding and information theory. In particular, factor-graph (FG) based modeling of codes [1] has facilitated efficient message-passing algorithms in decoding capacity-achieving codes, and normal realizations of codes on graphs [2] have brought to surface an elegant duality property of codes in a graphical context. It is also remarkable that these advances in the information theory community have resonated with the recent development in the computer science and statistical physics communities, where joint efforts are made towards developing efficient algorithmic solvers for various intractable computational problems and towards characterizing the hardness of such problems. For example, Valiant [3] developed what he calls “holographic algorithms” as polynomial-time solvers for a number of problem families which were previously unknown to be in P. In a related work [4], Chernyak and Chertkov characterized a rich family of binary graphical models on planar graphs for which the computation of the partition function is easy. Valiant’s holographic algorithms have also recently found applications in the information theory community for solving certain constrained-coding capacity problems [5].

Inspired by the holographic algorithms and making a joint effort with Forney [6], the authors of [7] have recently re-formalized the notion of normal factor graphs (NFGs) by introducing the “exterior function” semantics, which unifies Forney’s duality result in normal realizations [2] and the so-called “holant theorem” used by Valiant [3] to establish the holographic algorithms. This has ignited, most recently, the work of Forney and Vontobel [8], which pushes further the power of NFGs as an elegant graphical model.

In this paper, we point out yet another direction that NFGs may demonstrate their advantages, namely, their use as a linear algebraic tool. Much of this work is motivated by the notion of “trace diagrams,” which has been recently recognized in the mathematics community as an elegant and intuitive “diagrammatic approach” towards expressing notions in linear algebra [9], [10], [11], [12], [13]. Our main objective is to demonstrate

how NFGs can be used to provide diagrammatic proofs of some results, and also to show their ability to generalize trace diagrams. Many of the identities used as examples are reproductions of results from [13] in the language of NFGs and we make no claim of the originality of these results. However, we believe that the NFG approach is more general than trace diagrams and may potentially provide a wider range of algebraic tools. Indeed, the lucid approach of NFGs enables us to prove a conjecture of Peterson on the Pfaffian [14] [12].

The remainder of this paper is structured as follows. Section II is primarily a review of the NFG framework under the exterior-function semantics. Section III introduces some additional notions of NFGs which will be useful in deriving the results of this paper. The main results and example identities are presented in Section IV. The paper is concluded in Section V.

## II. PRELIMINARIES

### A. Some Notation

In this subsection, we briefly list some of our notations. All functions have finite domains (which might be different) and assume their values from the same field  $\mathbb{F}$ . A univariate function is written as  $f(x)$ , which will also denote the  $x$ th component of a vector  $f$ . That is,  $f$  will be used to denote a univariate function or a vector, and the distinction will be clear from the context – all vectors in this work are column vectors. Similarly, a bivariate function is denoted  $A(x_1, x_2)$ , which will also be used to denote the  $(x_1, x_2)$ th element of a matrix  $A$ . For any matrix  $A$ , we will use  $A^T$  to denote its transpose, and if  $A$  is a square matrix, we use  $\det(A)$  and  $\text{tr}(A)$  to denote its determinant and trace, respectively.

For any positive integer  $n$ , let  $N = \{1, \dots, n\}$  and let  $S_n$  be the symmetric group on  $N$ . A permutation  $\sigma \in S_n$  such that  $\sigma(j) = i_j$  for all  $j \in N$  will be written as  $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ . Finally, we use  $\varepsilon$  to denote the Levi-Civita symbol, which is defined as

$$\varepsilon(x_1, \dots, x_n) = \begin{cases} \text{sgn} \begin{pmatrix} 1 & \cdots & n \\ x_1 & \cdots & x_n \end{pmatrix}, & \begin{pmatrix} 1 & \cdots & n \\ x_1 & \cdots & x_n \end{pmatrix} \in S_n \\ 0, & \text{otherwise} \end{cases}$$

for all  $(x_1, \dots, x_n) \in N^n$ . Note that the domain of  $\varepsilon$  is specified by the number of its arguments and hence we need not be explicit about it.

### B. Normal Factor Graphs

Let  $V$ ,  $E$ , and  $D$  be disjoint finite sets and  $\{\mathcal{X}_i : i \in E \cup D\}$  be a collection of finite alphabets. Further, for any  $\mathcal{I} \subseteq E \cup D$ , we use  $x_{\mathcal{I}}$  to denote the variables set  $\{x_i : i \in \mathcal{I}\}$  where  $x_i \in \mathcal{X}_i$  for all  $i \in \mathcal{I}$ . A *normal factor graph* (NFG) [2] is a quadruple  $\Omega = (V, E, D, f_V)$  defined as follows.

- $(V, E)$  is a graph with vertex set  $V$  and edge set  $E$  where an edge is incident on exactly two vertices.
- $D$  is a set (possibly empty) of *dangling edges* where in contrast to a regular edge, a dangling edge is incident on *exactly one* vertex in  $V$ .
- Using  $\mathcal{E}(v) = \{e \in E \cup D : e \text{ is incident on } v\}$  to denote the set of edges and dangling edges incident on the vertex  $v \in V$ , each vertex  $v \in V$  is associated with a *local function*  $f_v$  involving precisely the variables set  $x_{\mathcal{E}(v)}$ , and  $f_V$  is the collection of all local functions, namely, the set  $\{f_v : v \in V\}$ .

Given an NFG  $\mathcal{G} = (V, E, D, f_V)$ , we may associate with it the unique form

$$\sum_{x_E} \prod_{v \in V} f_v(x_{\mathcal{E}(v)}).$$

We refer to this as the *sum-of-products form represented by  $\mathcal{G}$* . Clearly, the sum-of-products form is a function of  $x_D$ , which we refer to as the *exterior function realized by  $\mathcal{G}$*  and denote it by  $Z_{\mathcal{G}}(x_D)$ . It is not hard to see that while a sum-of-products form is represented by a unique NFG, an exterior function may be realized by an infinite class of NFGs. Finally, we say that two NFGs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  are *equivalent*, and write  $\mathcal{G}_1 = \mathcal{G}_2$ , if  $Z_{\mathcal{G}_1} = Z_{\mathcal{G}_2}$ . In some subsequent figures we may equate an NFG  $\mathcal{G}$  to a function  $f$ , which strictly speaking is not correct, but what we really mean is that  $Z_{\mathcal{G}} = f$ . Such practice is unlikely to raise confusion and makes some of the analysis more transparent.

A particular sum-of-products form that will be useful in developing some of the tools in this paper is what we refer to as the “simple” sum-of-products form. We define the *simple* sum-of-products form, denoted  $\langle \cdot | \cdot \rangle$ , as the sum-of-products form involving exactly two functions with the summation being over all common variables. More explicitly, for any functions  $f(x_{\mathcal{I}})$  and  $g(x_{\mathcal{J}})$ , the simple sum-of-products form is defined as

$$\langle f | g \rangle = \sum_{x_{\mathcal{I} \cap \mathcal{J}}} f(x_{\mathcal{I}}) g(x_{\mathcal{J}}).$$

Fig. 1 shows the NFG  $\mathcal{G}$  representing  $\langle f | g \rangle$ . In this figure each set of variables is treated as a single variable and hence represented by a single edge. If such a set is empty, the corresponding edge (or dangling edge) simply disappears from the figure. With this understanding, the NFG in Fig. 1 may encode several notions from linear algebra, namely:

- $\mathcal{I} \cap \mathcal{J} = \emptyset$ , then  $\mathcal{G}$  corresponds to the tensor product. If further,
  - $\mathcal{I} = \emptyset$  and  $\mathcal{J} \neq \emptyset$ , then the tensor product becomes the scalar-vector product.
  - $\mathcal{I} = \mathcal{J} = \emptyset$ , then the tensor product becomes the scalar-scalar product.
- $\mathcal{I} \cap \mathcal{J}$ ,  $\mathcal{I} \setminus \mathcal{J}$ , and  $\mathcal{J} \setminus \mathcal{I}$  are all non-empty, then  $\mathcal{G}$  corresponds to the matrix-matrix product.

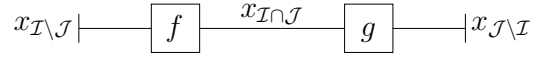


Fig. 1: The NFG  $\mathcal{G}$  representing the simple sum-of-products form  $\langle f | g \rangle$  where  $\mathcal{I} \setminus \mathcal{J} = \{i \in \mathcal{I} : i \notin \mathcal{J}\}$  and  $\mathcal{J} \setminus \mathcal{I} = \{j \in \mathcal{J} : j \notin \mathcal{I}\}$ .

- $\mathcal{I} \cap \mathcal{J} \neq \emptyset$ ,  $\mathcal{I} \setminus \mathcal{J} \neq \emptyset$ , and  $\mathcal{J} \setminus \mathcal{I} = \emptyset$ , then  $\mathcal{G}$  corresponds to the matrix-vector product.
- $\mathcal{I} = \mathcal{J} \neq \emptyset$  (i.e.,  $\mathcal{I} \cap \mathcal{J} \neq \emptyset$  and  $\mathcal{I} \setminus \mathcal{J} = \mathcal{J} \setminus \mathcal{I} = \emptyset$ ), then  $\mathcal{G}$  corresponds to the vector dot product.

In this work we find it helpful to adopt the notion of “ciliation” [13] to explicitly indicate the local orderings of variables at each node. To this end, we add a dot on each node to mark the edge carrying the local function’s first argument and assume the rest of arguments are encountered in a counter-clockwise manner, cf. Fig. 2. We will not insist on ciliations in every occasion and use them only to facilitate the analysis.

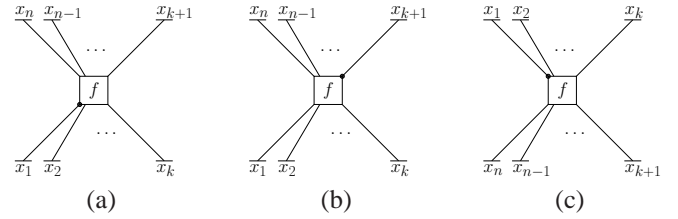


Fig. 2: Three assignments of ciliations illustrating different orderings of arguments: (a)  $f(x_1, \dots, x_n)$ , (b)  $f(x_{k+1}, \dots, x_n, x_1, \dots, x_k)$  and (c)  $f(x_1, x_n, x_{n-1}, \dots, x_2)$ .

As an example of the effect of different ciliation arrangements on the realized exterior function, consider the NFGs in Fig. 3 where  $A$  and  $B$  are  $n \times n$  matrices. Then, it is easy to verify that  $Z_{\mathcal{G}_1} = AB$ ,  $Z_{\mathcal{G}_2} = AB^T$ ,  $Z_{\mathcal{G}_3} = A^T B^T$ , and  $Z_{\mathcal{G}_4} = A^T B$ , when  $Z_{\mathcal{G}_i}$ ,  $i = 1, \dots, 4$ , is viewed as a matrix with rows and columns indexed by  $x_1$  and  $x_2$ , respectively, and the product is the regular matrix product.

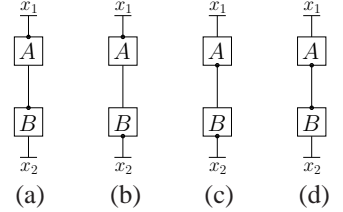


Fig. 3: Different arrangements of ciliations: (a)  $\mathcal{G}_1$ , (b)  $\mathcal{G}_2$ , (c)  $\mathcal{G}_3$ , and (d)  $\mathcal{G}_4$ .

We end this subsection with an example. Consider the NFG  $\mathcal{G}$  in Fig. 4 (a) where  $A$  is an  $n \times n$  matrix. Then from the definition of the exterior function, we have

$$Z_{\mathcal{G}} = \sum_i A(i, i) = \text{tr}(A).$$

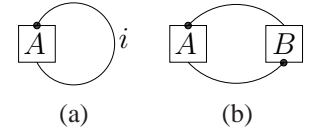


Fig. 4: (a) An NFG that realizes  $\text{tr}(A)$  and (b) an NFG that illustrates  $\text{tr}(AB) = \text{tr}(BA)$ .

This is the reason behind the name “trace diagrams” in [9], [10], [11], [12], [13]. Briefly, a *trace diagram* is an NFG where each node has degree one, degree two, or is associated a Levi-Civita symbol. As another example, let  $A$  be an  $m \times n$  matrix and  $B$  be an  $n \times m$  matrix. Then traversing the NFG in Fig. 4 (b) counter-clockwise starting from the upper edge and then starting from the lower edge illustrates the well-known identity,  $\text{tr}(AB) = \text{tr}(BA)$ .

### C. Vertex Grouping and Vertex Splitting

Given an NFG  $\mathcal{G}$ , we may group any two vertex functions of  $\mathcal{G}$ , say  $f$  and  $g$ , and replace them with a single function node realized by  $\langle f|g \rangle$ . We refer to this process as *vertex grouping* – also known as “closing the box” [15]. Conversely, if  $h$  is a vertex node of  $\mathcal{G}$  that is realized by  $\langle f|g \rangle$  for some  $f$  and  $g$ , then we may split the vertex node  $h$  into two function nodes  $f$  and  $g$ . We refer to this as *vertex splitting* – also known as “opening the box” [15]. A graphical illustration of vertex grouping and vertex splitting is shown in Fig. 5. Note that when we put a dashed box around  $f$  and  $g$  we mean they are replaced with the single function node  $\langle f|g \rangle$ , in other words, Figs. 5 (b) and (c) refer to the same NFG.

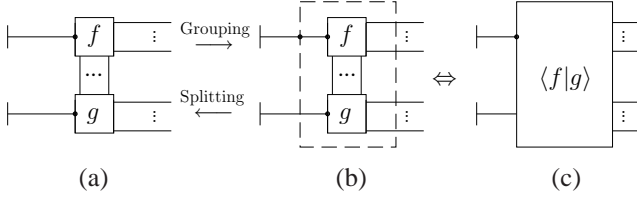


Fig. 5: Vertex grouping and vertex splitting

Since the range of all the functions in an NFG is a field, the following lemma, which will be the main proving tool to most of the subsequent results, is clear<sup>1</sup>.

**Lemma 1.** *The exterior function realized by any NFG is invariant under vertex grouping and vertex splitting.*

Clearly one may group an arbitrary number of nodes in an NFG by recursively grouping pairs of vertices; ultimately, one may replace an NFG  $\mathcal{G}$  with a single function node realizing the exterior function  $Z_{\mathcal{G}}$ .

### III. DERIVED NFGs: SCALING, ADDITION, AND SUBTRACTION OF NFGs

We have already seen vertex grouping and splitting as a means of manipulating NFGs. Now, given two NFGs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  with disjoint dangling edge sets, one of the most basic ways to obtain a new NFG, say  $\mathcal{G}$ , from  $\mathcal{G}_1$  and  $\mathcal{G}_2$  is to stack  $\mathcal{G}_1$  and  $\mathcal{G}_2$  beside each other. Grouping the vertices of  $\mathcal{G}_1$  and  $\mathcal{G}_2$ , it is clear that  $Z_{\mathcal{G}}$  is the tensor product of  $Z_{\mathcal{G}_1}$  and  $Z_{\mathcal{G}_2}$ . Of particular interest is the case when  $\mathcal{G}_2$  is a single node of degree zero. In this case  $Z_{\mathcal{G}_2}$  is a constant, say  $\lambda$ , and so  $Z_{\mathcal{G}} = \lambda Z_{\mathcal{G}_1}$ . Moreover, we write  $\mathcal{G} = \lambda \mathcal{G}_1$  and say  $\mathcal{G}$  is a *scaled* version of  $\mathcal{G}_1$  with  $\lambda$  being the *scaling* factor. Graphically, we show scaling by putting the scaling factor beside the NFG.<sup>2</sup>

Given two NFGs  $\mathcal{G}_1$  and  $\mathcal{G}_2$  with the same set of dangling edges, we may graphically construct a *compound* NFG, denoted  $\mathcal{G}_1 + \mathcal{G}_2$ , by drawing the two graphs with the plus sign, “+”, between them, Fig. 6. We use this compound NFG to represent the function  $Z_{\mathcal{G}_1} + Z_{\mathcal{G}_2}$ , and more formally, we say the compound NFG realizes  $Z_{\mathcal{G}_1} + Z_{\mathcal{G}_2}$ . Finally, the “subtraction of  $\mathcal{G}_2$  from  $\mathcal{G}_1$ ” compound NFG, denoted  $\mathcal{G}_1 - \mathcal{G}_2$ , is defined as  $\mathcal{G}_1 + (-1)\mathcal{G}_2$ .

<sup>1</sup> We remark that all is needed for Lemma 1 to hold is the distributive law, and so it is still true even when the range of the local functions of the NFG is a semi-ring rather than a field.

<sup>2</sup> An equivalent way to obtain  $\mathcal{G}$  directly from  $\mathcal{G}_1$  is to replace any function node, say  $f$ , in  $\mathcal{G}_1$  with the function node  $\lambda f$  in  $\mathcal{G}$ , and to leave everything else the same.

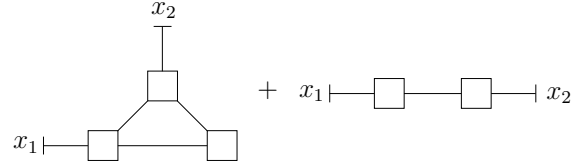


Fig. 6: Addition of two NFGs.

### IV. SOME LINEAR ALGEBRA USING NFGs

The main purpose of this section is to demonstrate how NFGs can be used to establish various identities from linear algebra in a simple and intuitive diagrammatic manner. Throughout the section, we will liberally make use of the following lemma about the Levi-Civita symbol, which easily follows from its definition.

**Lemma 2.** *For any positive integer  $n$ , it holds that*

$$\varepsilon(x_1, x_2, \dots, x_{n-1}, x_n) = (-1)^{n-1} \varepsilon(x_2, x_3, \dots, x_n, x_1),$$

and so, when  $n$  is odd,  $\varepsilon$  becomes invariant under cyclic-shifts of its arguments.

#### A. Kronecker Delta

The Kronecker delta, denoted  $\delta(x)$ , is the indicator function evaluating to one if and only if  $x = 0$ . An NFG realizing  $\delta(x - y)$  is shown in Fig. 7 (a), which from now on will be abbreviated as the NFG in (b).

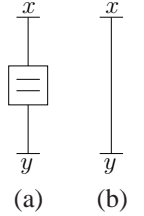


Fig. 7: Both NFGs in (a) and (b) realize  $\delta(x - y)$ .

#### B. Cross Product

In this subsection we look at the cross product of vectors of length 3. Let  $\mathcal{G}$  be as in Fig. 8, then direct computation shows that

$$Z_{\mathcal{G}}(1) = u(2)v(3) - u(3)v(2),$$

$$Z_{\mathcal{G}}(2) = u(3)v(1) - u(1)v(3),$$

$$Z_{\mathcal{G}}(3) = u(1)v(2) - u(2)v(1).$$

I.e., when viewed as vectors in  $\mathbb{F}^3$ , the exterior function  $Z_{\mathcal{G}}$  is the cross product  $u \times v$ .

A well-known fact, and straightforward to prove, is the following identity about the contraction of two Levi-Civita symbols, namely

$$\sum_t \varepsilon(y_1, y_2, t) \varepsilon(t, x_2, x_1) = \delta(x_1 - y_2) \delta(x_2 - y_1) - \delta(x_1 - y_1) \delta(x_2 - y_2).$$

A graphical illustration of the Levi-Civita contraction identity is shown in Fig. 9.

Gluing single variate nodes  $w$ ,  $s$ ,  $u$ , and  $v$ , respectively, on the dangling edges  $x_1$ ,  $x_2$ ,  $y_1$ , and  $y_2$  in Fig. 9 results in Fig. 10. From the previous analysis (in particular, Lemmas 1 and 2), the following identities are then clear:

$$\begin{aligned} (u \times v) \cdot (s \times w) &= ((u \times v) \times s) \cdot w = (w \times (u \times v)) \cdot s \\ &= ((s \times w) \times u) \cdot v = (v \times (s \times w)) \cdot u \\ &= (u \cdot s)(v \cdot w) - (u \cdot w)(v \cdot s). \end{aligned}$$

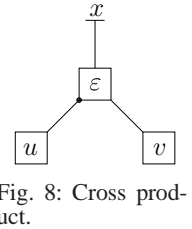


Fig. 8: Cross product.

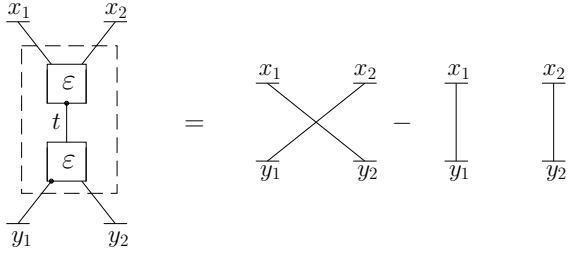


Fig. 9: Contraction of two Levi-Civita symbols.

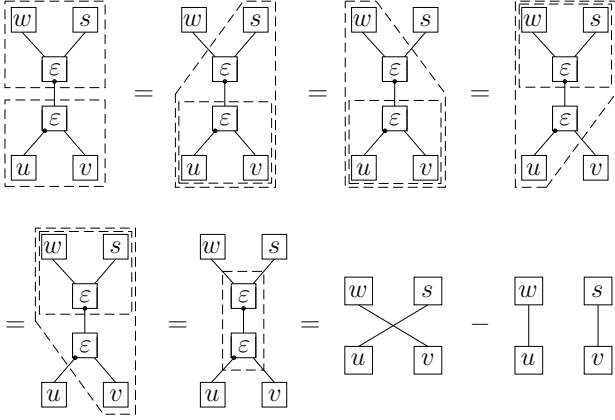


Fig. 10: A proof of  $(u \times v) \cdot (s \times w) = \dots = (u \cdot s)(v \cdot w) - (u \cdot w)(v \cdot s)$ .

Many identities regarding the cross product may be obtained in the same way. For instance, let  $m_a$ ,  $m_b$ ,  $m_c$ , and  $m_d$  be arbitrary positive integers and let  $\{a_i : 1 \leq i \leq m_a\}$ ,  $\{b_i : 1 \leq i \leq m_b\}$ ,  $\{c_i : 1 \leq i \leq m_c\}$ , and  $\{d_i : 1 \leq i \leq m_d\}$  be four collections of length-three vectors. Further, let  $A$  be the matrix whose  $i$ th column is  $a_i$ , i.e.,  $A = (a_i : 1 \leq i \leq m_a)$ , and similarly let  $B = (b_i : 1 \leq i \leq m_b)$ ,  $C = (c_i : 1 \leq i \leq m_c)$ , and  $D = (d_i : 1 \leq i \leq m_d)$ . We claim that if  $m_a = m_d = m$  and  $m_b = m_c = m'$  for some  $m$  and  $m'$ , then

$$\sum_{i=1}^m \sum_{j=1}^{m'} (a_i \times b_j) \cdot (c_j \times d_i) = \text{tr}(AD^T BC^T) - \text{tr}(BC^T) \text{tr}(AD^T).$$

A detailed proof of this identity is shown in Fig. 11.

We emphasize that in practice, one does not need to go through all the steps in Fig. 11, i.e., one should be able to deduce the identity easily from the last two NFGs. Slight variations of the NFG before the last equivalence in Fig. 11 may be used to prove more identities about the cross product. For instance, when  $m_a = m_b = m$  and  $m_c = m_d = m'$  for some  $m$  and  $m'$ , Fig. 12 (a) shows that

$$\sum_{i=1}^m \sum_{j=1}^{m'} (a_i \times b_i) \cdot (c_j \times d_j) = \text{tr}(AB^T DC^T) - \text{tr}(AB^T CD^T),$$

and when  $m_a = 1$  and  $m_b = m_c = m$  for some  $m$ , Fig. 12 (b) proves  $\sum_{i=1}^m (a_1 \times b_i) \times c_i = (BC^T) a_1 - \text{tr}(BC^T) a_1$ .

We invite the reader to verify some of the identities above using traditional methods and contrast with the current diagrammatic approach.

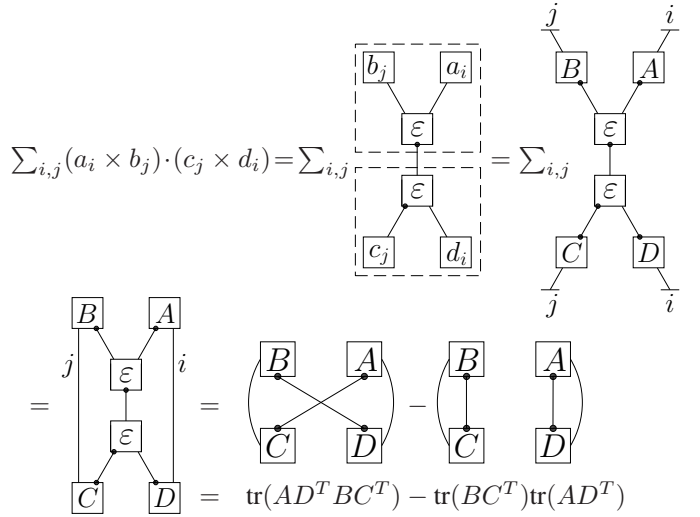


Fig. 11

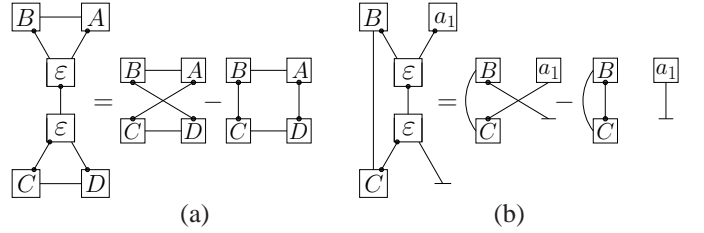


Fig. 12

### C. Determinant

For any  $n \times n$  matrix  $A$ , the determinant is defined as

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{j=1}^n A(j, \sigma(j)).$$

From the definitions of exterior function and determinant, it follows that

$$\begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ A \quad \dots \quad A \end{array} = \det(A) \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \dots \end{array}$$

From this, several properties of the determinant may be obtained, for instance

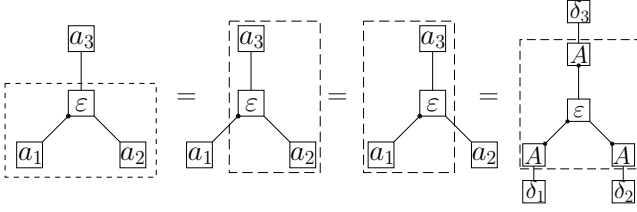
$$\begin{aligned} \det(AB) &= \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ AB \quad \dots \quad AB \end{array} = \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ A \quad \dots \quad A \end{array} \quad \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ B \quad \dots \quad B \end{array} \\ = \det(A) \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ B \quad \dots \quad B \end{array} = \det(A) \det(B) \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \dots \end{array} \end{aligned}$$

shows that  $\det(AB) = \det(A) \det(B)$  for any  $n \times n$  matrices  $A$  and  $B$ . Another example follows from

$$\begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \epsilon \quad \dots \quad \epsilon \end{array} = \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ A^T \quad \dots \quad A^T \end{array} \quad \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \epsilon \quad \dots \quad \epsilon \end{array} \\ = \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ A \quad \dots \quad A \end{array} \quad \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \epsilon \quad \dots \quad \epsilon \end{array} \\ = \det(A) \begin{array}{c} \epsilon \\ \swarrow \quad \searrow \\ \epsilon \quad \dots \quad \epsilon \end{array} \end{array}$$

showing  $\det(A^T) = \det(A)$ . As a final example, we have





which proves

$$(a_1 \times a_2) \cdot a_3 = (a_2 \times a_3) \cdot a_1 = (a_3 \times a_1) \cdot a_2 = \det(A),$$

where  $A$  is the  $3 \times 3$  matrix  $(a_1 \ a_2 \ a_3)$  and for each  $i \in \{1, 2, 3\}$ ,  $\delta_i(x) = \delta(x - i)$  for all  $x \in \{1, 2, 3\}$ .

### D. Pfaffian

Given a  $2n \times 2n$  skew-symmetric matrix  $A$ , the *Pfaffian* of  $A$ , denoted  $\text{Pf}(A)$ , is defined as

$$\text{Pf}(A) = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \prod_{i=1}^n A(\sigma(2i-1), \sigma(2i)).$$

Before we proceed, we need the following lemma.

**Lemma 3.** *For any positive integer  $n$ , let  $\tau \in S_{2n}$  be such that  $\tau(2k-1) = k$  and  $\tau(2k) = 2n - (k-1)$ ,  $1 \leq k \leq n$ . Then  $\tau$  is an even permutation.*

*Proof:* For brevity, we use the tuple  $(i_1, \dots, i_{2n})$  to refer to the permutation  $\begin{pmatrix} 1 & \dots & 2n \\ i_1 & \dots & i_{2n} \end{pmatrix}$ . Now, given the identity permutation  $(1, \dots, n, n+1, \dots, 2n)$ , it is possible to reflect the second half of the permutation so that it becomes  $(1, \dots, n, 2n, \dots, n+1)$ ; this clearly can be done using  $\lfloor \frac{n}{2} \rfloor$  swaps. Next, we interleave the two halves of the  $2n$ -tuple by performing the following  $n-1$  steps: Starting with  $k=1$  and ending with  $k=n-1$ , at step  $k$  imagine a window that starts just before the  $2k$ th element of the tuple and ends just after the  $n+k$ th element (i.e., it covers  $n-k+1$  elements). For the elements of the tuple covered by the window, we perform a cyclic-shift by one position to the right. It is easy to check that after the  $(n-1)$ th step we arrive to our desired permutation. Clearly, a one-position cyclic shift on  $m$  elements can be accomplished using  $m-1$  swaps. Hence, interleaving the tuple requires  $\sum_{m=2}^n (m-1) = \frac{n(n-1)}{2}$  swaps. Therefore, in total our permutation can be written in terms of  $\lfloor \frac{n}{2} \rfloor + \frac{n(n-1)}{2}$  swaps and the claim follows by noting that this number is even for any positive integer  $n$ . ■

The following proposition affirmatively proves Peterson's conjecture on the Pfaffian [14] [12].

**Proposition 1.** *Let  $A$  be a  $2n \times 2n$  skew-symmetric matrix and let  $\mathcal{G}$  be as in Fig. 13. Then  $Z_{\mathcal{G}} = n! \cdot 2^n \cdot \text{Pf}(A)$ .*

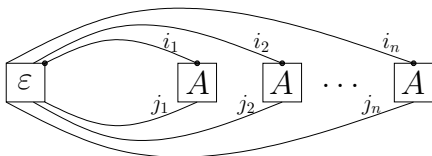


Fig. 13: The NFG  $\mathcal{G}$  from Proposition 1.

*Proof:* We have

$$\begin{aligned} Z_{\mathcal{G}} &= \sum_{(i_1, \dots, i_n, j_n, \dots, j_1) \in N^{2n}} \varepsilon(i_1, \dots, i_n, j_n, \dots, j_1) \prod_{k=1}^n A(i_k, j_k) \\ &= \sum_{\sigma \in S_{2n}} \text{sgn}(\sigma) \prod_{k=1}^n A(\sigma(k), \sigma(2n - (k-1))). \end{aligned}$$

Now let  $\sigma = \begin{pmatrix} 1 & 2 & \dots & n & n+1 & \dots & 2n \\ i_1 & i_2 & \dots & i_n & j_n & \dots & j_1 \end{pmatrix}$  and  $\sigma' = \begin{pmatrix} 1 & 2 & 3 & 4 & \dots & 2n-1 & 2n \\ i_1 & j_1 & i_2 & j_2 & \dots & i_n & j_n \end{pmatrix}$  and note that  $\sigma' = \sigma \circ \tau$  where  $\tau$  is as in Lemma 3. Hence,  $\text{sgn}(\sigma') = \text{sgn}(\sigma) \cdot \text{sgn}(\tau) = \text{sgn}(\sigma)$ , where the last equality follows from Lemma 3. Further, note that for all  $k \in \{1, \dots, n\}$  we have  $\sigma(k) = i_k = \sigma'(2k-1)$  and  $\sigma(2n - (k-1)) = j_k = \sigma'(2k)$ . Finally, it is clear that as  $\sigma$  runs over  $S_{2n}$ ,  $\sigma'$  runs over  $S_{2n}$ . Hence,  $Z_{\mathcal{G}} = \sum_{\sigma' \in S_{2n}} \text{sgn}(\sigma') \prod_{k=1}^n A(\sigma'(2k-1), \sigma'(2k))$ , and the claim follows by the definition of the Pfaffian. ■

### V. CONCLUSION

This paper presents NFGs as an intuitive tool that facilitates proofs and enhances understanding of some notions and identities in linear algebra. We feel that this new look at NFGs, together with the related work [8], may result in a better understanding of NFGs as a model of codes and provide deeper understanding of the message-passing algorithms. Further, it is our hope that the use of NFGs as an algebraic tool may open doors for developing new algorithmic tools that are useful in a broader range of disciplines.

### ACKNOWLEDGMENT

The authors wish to thank David Forney for several discussions on the subject.

### REFERENCES

- [1] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [2] G. D. Forney, Jr., "Codes on graphs: Normal realizations," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 520–548, 2001.
- [3] L. G. Valiant, "Holographic algorithms (extended abstract)," in *Proc. 45th Annual IEEE Symp. on Foundations of Computer Science*, Rome, Italy, Oct. 2004, pp. 306–315.
- [4] V. Y. Chernyak and M. Chertkov, "Planar graphical models which are easy," *J. Stat. Mech.*, P11007, Nov. 2010.
- [5] M. Schwartz and J. Bruck, "Constrained codes as networks of relations," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 2179–2195, 2008.
- [6] G. D. Forney, Jr., "Codes on graphs: Duality and MacWilliams identities," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1382–1397, 2011.
- [7] A. Al-Bashabsheh and Y. Mao, "Normal factor graphs and holographic transformations," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 752–763, Feb. 2011.
- [8] G. D. Forney, Jr. and P. O. Vontobel, "Partition functions of normal factor graphs," in *Proc. Information Theory and Applications Workshop*, San Diego, CA, Feb. 2011.
- [9] P. Cvitanović, *Group Theory: Birdtracks, Lie's, and Exceptional Groups*. Princeton Univ. Press, 2008.
- [10] R. Penrose, *Combinatorial Mathematics and Its Applications*, chapter Applications of negative dimensional tensors, D. Welsh, Ed. Academic Press, 1971.
- [11] —, *The Road to Reality: A Complete Guide to the Laws of the Universe*. Jonathan Cape, London, 2004.
- [12] E. Peterson, "Unshackling linear algebra from linear notation," 2009, available at arXiv: 0910.1362.
- [13] S. Morse and E. Peterson, "Trace diagrams, signed graph coloring and matrix minors," *Involve, a Journal of Mathematics*, vol. 3, no. 1, 2010.
- [14] E. Peterson, "On a diagrammatic proof of the Cayley-Hamilton theorem," 2009, available at arXiv:0907.2364v1.
- [15] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Sig. Proc. Mag.*, vol. 21, no. 1, pp. 24–41, 2004.